

BOJANA V. ROSIĆ, ALEXANDER LITVINENKO,
OLIVER PAJONK AND HERMANN G. MATTHIES

DIRECT BAYESIAN UPDATE OF POLYNOMIAL CHAOS REPRESENTATIONS

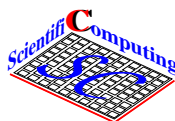


INSTITUTE OF SCIENTIFIC COMPUTING
CARL-FRIEDRICH-GAUSS-FAKULTÄT
TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG

Brunswick, Germany

This document was created February 2011 using L^AT_EX 2_ε.

Informatikbericht 2011-02
Institute of Scientific Computing
Technische Universität Braunschweig
Hans-Sommer-Straße 65
D-38106 Braunschweig, Germany



url: www.wire.tu-bs.de
mail: wire@tu-bs.de

Copyright © by Bojana V. Rosić, Alexander Litvinenko, Oliver Pajonk
and Hermann G. Matthies

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted in connection with reviews or scholarly analysis. Permission for use must always be obtained from the copyright holder.

Alle Rechte vorbehalten, auch das des auszugsweisen Nachdrucks, der auszugsweisen oder vollständigen Wiedergabe (Photographie, Mikroskopie), der Speicherung in Datenverarbeitungsanlagen und das der Übersetzung.

Direct Bayesian update of polynomial chaos representations

Bojana V. Rosić, Alexander Litvinenko, Oliver Pajonk and
Hermann G. Matthies

Institute of Scientific Computing, TU Braunschweig
wire@tu-bs.de

Abstract

We present a fully deterministic approach to a probabilistic interpretation of inverse problems in which unknown quantities are represented by random fields or processes, described by a non-Gaussian prior distribution. The description of the introduced random fields is given in a “white noise” framework, which enables us to solve the stochastic forward problem through Galerkin projection onto polynomial chaos. With the help of such representation, the probabilistic identification problem is cast in a polynomial chaos expansion setting and the linear Bayesian form of updating. By introducing the Hermite algebra this becomes a direct, purely algebraic way of computing the posterior, which is inexpensive to evaluate. In addition, we show that the well-known Kalman filter method is the low order part of this update. The proposed method has been tested on a stationary diffusion equation with prescribed source terms, characterised by an uncertain conductivity parameter which is then identified from limited and noisy data obtained by a measurement of the diffusing quantity.

Keywords: minimum squared error estimate, minimum variance estimate, polynomial chaos expansion, linear Bayesian update, Kalman filter

1 Introduction

The mathematical model of a physical system is often characterised by quantities which may be described as uncertain due to incomplete knowledge.

The reduction of the uncertainty with the help of noisy and incomplete data, obtained by measuring a function of the system response due to different excitations, is the subject of this paper. This is a so-called inverse problem of identification. The goal is to circumvent the “ill-posedness” of the inverse problem in the sense of Hadamard by a Bayesian regularisation method [49, 48], which uses a priori knowledge as additional information to the given set of data. In this paper we especially focus on the use of a linear Bayesian approach [18] in the framework of “white noise” analysis.

In order to fix ideas, a Darcy-like flow through the porous medium is considered as an example. In this particular case, the mathematical model of the system represents a finite-element discretisation of the governing stationary diffusion equation. The corresponding identification problem is cast in a linear Bayesian framework [2, 19, 21, 29, 35, 36, 50, 51], taking a probabilistic model for the uncertain conductivity field with the prior assumption in a form of a lognormal random field with some covariance function. The goal is to update the description of the conductivity field from the data, obtained by measuring the hydraulic head in certain locations due to applied hydraulic loading.

Since the parameters of the model to be estimated are uncertain, all relevant information may be obtained via their probability density functions. In order to extract information from the posterior density most estimates take the form of integrals over the posterior. These integrals may be numerically computed via asymptotic, deterministic or sampling methods. The most often used technique represents a Markov chain Monte Carlo (MCMC) method [13], which takes the posterior distribution for the asymptotic one. This then allows the approximation of the desired posterior expectations by ergodic averages. With the intention of accelerating the MCMC method many authors [29, 36, 45] have tried to introduce the stochastic spectral methods into the computation. Expanding the prior random process into the polynomial chaos [52] or a Karhunen-Loève (K-L) expansion [17, 37, 38], the inverse problem transforms to inference on a truncated sequence of weights of the KL modes. This approach employs the dimensionality reduction of the problem by estimating the likelihood function from the approximated solution and further sampling it in the MCMC way. Pence [45] proposed a new approach which combines polynomial chaos theory with maximum likelihood estimation, where the parameter estimates are calculated in a recursive or iterative manner. In order to improve the acceptance probability of proposed moves, Christen and Fox [10] have applied a local linearisation of the forward model, while some authors [7, 33, 31] have tried to employ the collocation methods as a more efficient sampling technique.

The previously mentioned methods require a large number of samples in order to obtain satisfying results. Due to this, they are not suitable for high dimensional problems and another procedure has to be introduced. The main idea of this work is to perform the Bayesian update of the polynomial chaos expansion of the a priori information to an a posteriori one without any sampling, but in a direct, purely algebraic way in a sequel to [44], where this kind of update was used in a case when the complete state of a dynamical system is measured. This idea has appeared independently in [8] in a simpler context.

The paper is organised in a following manner: Section 2 gives the mathematical description of an abstract physical system, whose identification is further considered in Section 3 in the framework of the introduced method. In Section 4 we describe the example scenario of the forward problem — the stationary diffusion equation with uncertainties. Section 5 outlines discretisation of the obtained equations in both the deterministic and stochastic context. Following this we are able to pose the inverse problem with respect to the polynomial chaos representation of quantities of consideration and propose the linear Bayesian estimator in Section 6. This allows us to use the Hermite algebra in the algorithm described in Section 7. The validation and testing of the proposed algorithm is shown in Section 8 for different cases of the analytical form of the “true” value of the model parameter. The paper is then concluded in Section 9.

2 Abstract Problem Setting

Let us represent the — so-called — forward model of some physical system by an operator A , which describes the relation between the model parameters q , external influence f , and the system state u :

$$A(u; q) = f. \quad (1)$$

For simplicity we assume the response u to be an element of a vector space \mathcal{U} , and the external influence in the dual space \mathcal{U}^* . More importantly here, we also assume the model parameters q to be elements of some vector space \mathcal{Q} ; and that Eq. (1) defines a well-posed problem such that there exists a unique “solution” u satisfying:

$$u = S(q; f), \quad (2)$$

where S is the solution operator describing the explicit relationship between u and the model parameters q . We further define an observation operator Y

relating the complete model response u to an observation y in some vector space \mathcal{Y}

$$y = Y(q; u) = Y(q; S(q; f)). \quad (3)$$

However, the measurements are in practice always disturbed by some kind of error ϵ , which determines the difference between the actual value \hat{y} of a measured quantity and the observed value z

$$z = \hat{y} + \epsilon. \quad (4)$$

The random elements in ϵ are assumed to be independent of the uncertainty in the model parameters q .

3 General Bayes Filter

3.1 The Identification or Assimilation Problem

Let us suppose that one may subject the system Eq. (1) to varying external influences f , and observe the output z , as the observation of the model parameters q is not possible. The goal is to update the knowledge of the possible parameter values q from observations z in a typical Bayesian setting, here chosen as an approach based on a minimisation of the squared error (MSE).

Let us assume that the true q is unknown, but as we are uncertain our best guess at the moment is a random variable q_f (the index f is for “forecast”) with values in \mathcal{Q} , i.e. $q_f : \Omega \rightarrow \mathcal{Q}$ is a measurable mapping, where Ω is the basic probability set of elementary events. The set Ω defines a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$, with \mathfrak{A} being a σ -algebra of subsets of Ω and \mathbb{P} a probability measure.

For simplicity we assume that \mathcal{Q} is a Hilbert space with inner product $\langle \cdot | \cdot \rangle_{\mathcal{Q}}$, and $\mathcal{S} := L_2(\Omega)$ the space of random variables of finite variance such that the \mathcal{Q} -valued random variables form a Hilbert space $\mathcal{Q} := \mathcal{Q} \otimes \mathcal{S} \cong L_2(\Omega, \mathcal{Q})$ with inner product

$$\langle\langle q_1 | q_2 \rangle\rangle_{\mathcal{Q}} := \mathbb{E} (\langle q_1(\omega) | q_2(\omega) \rangle_{\mathcal{Q}}). \quad (5)$$

A frequent procedure is then to take a closed subspace \mathcal{Q}_f of \mathcal{Q} as representing our current knowledge, and to assume that q_f is the orthogonal projection — i.e. the point closest to q in \mathcal{Q}_f in the norm generated by $\langle \cdot | \cdot \rangle_{\mathcal{Q}}$ — of q onto the subspace \mathcal{Q}_f .

In case one has no knowledge whatsoever prior to any observation, it is possible to take simply $q_f \equiv 0$ corresponding to $\mathcal{Q} = \{0\}$. Assume that from

the observations — which we assume to be in a Hilbert space $\mathcal{Y} := \mathcal{Y} \otimes \mathcal{S}$ of \mathcal{Y} -valued random variables — analogous to \mathcal{Q} — we obtain more information, and these observation random variables generate a subspace $\mathcal{Y}_0 \subseteq \mathcal{Y}$. On the other hand, from our prior knowledge q_f we *predict* the observation $y = Hq_f$, where H is a linear map from \mathcal{Q} onto \mathcal{Y} . Then these observations in \mathcal{Y}_0 generate the subspace

$$\mathcal{Q}_0 = H^* (\mathcal{Y}_0) \quad (6)$$

in \mathcal{Q} . The subspace $\mathcal{Q}_f + \mathcal{Q}_0 \subset \mathcal{Q}$ now represents the combined information both from our prior knowledge and from the observation. What we want now is the projection of q onto $\mathcal{Q}_f + \mathcal{Q}_0$ to assimilate the new observation, given that we know $q_f \in \mathcal{Q}_f$.

In this situation we may paraphrase a projection theorem from [32], a generalisation of the Gauss-Markov theorem, which in turn is the basis of most regression procedures:

Theorem 3.1. *In the setting just described, the random variable $q_a \in \mathcal{Q}$ — “a” stands for “assimilated” or “analysis” — is the orthogonal (MSE) projection of q onto the subspace $\mathcal{Q}_f + \mathcal{Q}_0$:*

$$q_a(\omega) = q_f(\omega) + K(z(\omega) - y(\omega)), \quad (7)$$

with q_f being the orthogonal projection onto \mathcal{Q}_f and K the “Kalman gain” operator given by

$$K = C_{q_f y} (C_y + C_\epsilon)^{-1}, \quad (8)$$

where $C_{q_f y} = \text{Cov}(q_f, y) = \mathbb{E}((q_f - \mathbb{E}(q_f)) \otimes (y - \mathbb{E}(y)))$, and similarly $C_y = \text{Cov}(y, y)$, $C_\epsilon = \text{Cov}(\epsilon, \epsilon)$.

In other words, $q_a(\omega)$ is the orthogonal projection of $q(\omega)$ onto $\mathcal{Q}_a = \mathcal{Q}_f + \mathcal{Q}_0$, which may be written as

$$\mathcal{Q}_a = \mathcal{Q}_f + \mathcal{Q}_0 = \mathcal{Q}_f \oplus \mathcal{Q}_i, \quad (9)$$

where the information gain (or innovation) space \mathcal{Q}_i is orthogonal to \mathcal{Q}_f . This orthogonal decomposition in Eq. (9) is reflected in Eq. (7), where $K(z(\omega) - y(\omega))$ generates \mathcal{Q}_i , and hence the variances of the terms in Eq. (7) are related through Pythagoras’s theorem.

Even if the spaces \mathcal{Q} and \mathcal{Y} were finite-dimensional, the spaces \mathcal{Q} and \mathcal{Y} are infinite-dimensional, and hence it is not possible to work with Theorem 3.1 directly in a computational setting. Hence, one approximates \mathcal{Q} by a finite dimensional subspace $\hat{\mathcal{Q}} := \mathcal{Q}_N \otimes \mathcal{S}_J$ with the orthogonal projector

$\hat{P} : \mathcal{Q} \rightarrow \hat{\mathcal{Q}}$, satisfying $\hat{P}^* = \hat{P}$. In other words, one “projects” the projection Eq. (7) onto the subspace $\hat{\mathcal{Q}}$:

$$\begin{aligned}\hat{q}_a(\omega) &= \hat{P}q_a(\omega) = \hat{P}(q_f(\omega) + K(z(\omega) - y(\omega))) = \hat{P}q_f(\omega) + \hat{P}K(z(\omega) - \hat{y}(\omega)) \\ &= \hat{q}_f(\omega) + \hat{P}K(z(\omega) - \hat{y}(\omega)),\end{aligned}\quad (10)$$

where $\hat{y}(\omega) = H\hat{P}q_f(\omega) = H\hat{q}_f(\omega)$. The orthogonal decomposition Eq. (9) is hence replaced by

$$\hat{P}\mathcal{Q}_a = \hat{\mathcal{Q}}_a = \hat{P}\mathcal{Q}_f \oplus \hat{P}\mathcal{Q}_i = \hat{\mathcal{Q}}_f \oplus \hat{\mathcal{Q}}_i. \quad (11)$$

With all spaces now finite-dimensional, the update in the finite-dimensional projection Eq. (10) reduces effectively to a matrix equation and is hence readily computed, see also Eq. (34), Eq. (38) and Eq. (39) in Section 6.1.

4 Stochastic Forward Problem — An Example

4.1 Stochastic Forward Problem

The particular model problem considered here is formally a stationary diffusion equation described by a conductivity parameter. It may, for example, describe the groundwater flow through a porous subsurface rock / sand formation [9, 15, 24, 41, 53]. Since the conductivity parameter in such cases is poorly known and may be considered as uncertain, one may model it as a random field and try to reconstruct it with data obtained by measurements.

Let us introduce a bounded spatial domain of interest $\mathcal{G} \subset \mathbb{R}^d$ together with the hydraulic head u , the conductivity parameter κ appearing in *Darcy’s* law for the seepage flow $q = -\kappa \nabla u$, and f as flow sinks and sources. For the sake of simplicity we only consider a scalar conductivity, although a conductivity tensor would be more appropriate. By applying the principle of conservation of mass one arrives at an equilibrium equation:

$$\begin{aligned}-\operatorname{div}(\kappa(x, \omega) \nabla u(x, \omega)) &= f(x, \omega) & \text{a.e. } x \in \mathcal{G}, \quad \mathcal{G} \subset \mathbb{R}^2, \\ u(x, \omega) &= 0 & \text{a.e. } x \in \partial\mathcal{G}.\end{aligned}\quad (12)$$

The conductivity κ and the source f are defined as random fields over the probability space Ω . Thus Eq. (12) is required to hold almost surely in ω , i.e. \mathbb{P} -almost everywhere.

As the conductivity κ has to be positive, and is thus restricted to a positive cone in a vector space, we consider its logarithm as the primary quantity, which may have any value. Assuming that it has finite variance one may

choose for maximum entropy a Gaussian distribution. Hence the conductivity is initially log-normally distributed. Such kind of assumption is known as *a priori* information/distribution:

$$\kappa(x) := \exp(q(x)), \quad q(x) \sim N(\mu_q, \sigma_q^2). \quad (13)$$

Also initially an exponential covariance function $\text{Cov}_q(x, y) = \sigma_q^2 \exp(-|x - y|/l_c)$ with prescribed covariance length l_c is chosen for $q(x)$.

In order to make sure that the numerical methods will work well, we strive to have similar overall properties of the stochastic system Eq. (12) as in the deterministic case (for fixed ω). For this to hold, it is necessary that the operator in Eq. (1) implicitly described by Eq. (12) is continuous and continuously invertible, i.e. both $\kappa(x, \omega)$ and $1/\kappa(x, \omega)$ have to be essentially bounded (have finite L_∞ norm) [5, 41, 38]:

$$\kappa(x, \omega) > 0 \quad \text{a.e.}, \quad \|\kappa\|_{L_\infty(\mathcal{G} \times \Omega)} < \infty, \quad \|1/\kappa\|_{L_\infty(\mathcal{G} \times \Omega)} < \infty. \quad (14)$$

Two remarks are in order here: one is that for a heterogeneous medium each realisation $\kappa(x, \omega)$ should be modelled as a tensor field. This would entail a bit more cumbersome notation and not help to explain the procedure any better. Hence for the sake of simplicity we stay with the unrealistically simple model of a scalar conductivity field. As a second remark we note that the conductivity has to be positive, and hence the set of possible conductivity fields is a cone — also a smooth manifold which can be equipped with a Lie group structure — in the vector space of random fields, but not a subspace. Therefore it is not possible to update it directly via a linear projection — see the discussion in [3, 4, 43] for symmetric positive matrices. Hence, we take the logarithm $q(x, \omega) = \log \kappa_f$, which maps the random field onto the tangent space at the neutral element of the mentioned Lie group. In this vector space one may use the projection setting of Theorem 4.1. After the update, the assimilated log of the random field q_a is transformed back to the posterior field κ_a by an exponential mapping. Due to this, the update formulas of Section 6 may be used on q .

The strong form given in Eq. (12) is not a good starting point for the Galerkin approach. Thus, as in the purely deterministic case, a variational formulation is needed, leading — via the *Lax-Milgram* lemma — to a well-posed problem. Hence, we search for $u \in \mathcal{U} := \mathcal{U} \otimes \mathcal{S}$ such that for all $v \in \mathcal{U}$ holds:

$$\mathbf{a}(v, u) := \mathbb{E}(\mathbf{a}(\omega)(v(\cdot, \omega), u(\cdot, \omega))) = \mathbb{E}(\langle \ell(\omega), v(\cdot, \omega) \rangle) =: \langle \ell, v \rangle. \quad (15)$$

Here $\mathbb{E}(b) := \mathbb{E}(b(\omega)) := \int_\Omega b(\omega) \mathbb{P}(d\omega)$ is the expected value of the random variable (RV) b . The double bracket $\langle \cdot, \cdot \rangle_{\mathcal{U}}$ is interpreted as duality pairing between \mathcal{U} and its dual space \mathcal{U}^* .

The bi-linear form \mathbf{a} in Eq. (15) is defined using the usual deterministic bi-linear (though parameter-dependent) form :

$$\mathbf{a}(\omega)(v, u) := \int_{\mathcal{G}} \nabla v(x) \cdot (\kappa(x, \omega) \nabla u(x)) \, dx, \quad (16)$$

for all $u, v \in \mathcal{U} := \dot{H}^1(\mathcal{G}) = \{u \in H^1(\mathcal{G}) \mid u = 0 \text{ on } \partial\mathcal{G}\}$. The linear form ℓ in Eq. (15) is similarly defined through its deterministic but parameter-dependent counterpart:

$$\langle \ell(\omega), v \rangle := \int_{\mathcal{G}} v(x) f(x, \omega) \, dx, \quad \forall v \in \mathcal{U}, \quad (17)$$

where f has to be chosen such that $\ell(\omega)$ is continuous on \mathcal{U} and the linear form ℓ is continuous on \mathcal{U} , the Hilbert space tensor product of \mathcal{U} and \mathcal{S} .

Let us remark that — loosely speaking — the stochastic weak formulation is just the expected value of its deterministic counterpart, formulated on the Hilbert tensor product space $\mathcal{U} \otimes \mathcal{S}$, i.e. the space of \mathcal{U} -valued RVs with finite variance, which is isomorphic to $L_2(\Omega, \mathbb{P}; \mathcal{U})$. In this way the stochastic problem can have the same theoretical properties as the underlying deterministic one, which is highly desirable for any further numerical approximation.

4.2 Measurement of Data

The coercive bilinear form in Eq. (15) defines a selfadjoint positive and continuous linear map

$$A : \mathcal{U} \mapsto \mathcal{U}^*, \quad (18)$$

which depends continuously on κ and hence on q . This operator is continuously invertible, and thus it defines the solution operator as it is given in Eq. (2), which also depends continuously on the model parameter q .

The goal of this paper is to identify the parameter κ from observations and a priori information. However, one cannot measure the conductivity $\kappa = \exp(q)$ directly — only some functional of the solution u ; here denoted by y . Let us assume that we perform the experiments providing us with measurements of y , made in finitely many patches L :

$$\hat{\mathcal{G}} := \{x_1, \dots, x_L\} \subset \mathcal{G}, \quad L := |\hat{\mathcal{G}}|. \quad (19)$$

An example of such a functional is the average hydraulic head:

$$y(u, \omega) := [\dots, y(x_j), \dots] \in \mathbb{R}^L, \quad y(x_j) = \int_{\mathcal{G}_j} u(x, \omega) dx, \quad (20)$$

where $\mathcal{G}_j \subset \mathcal{G}$ is a little patch centred at $x_j \in \hat{\mathcal{G}}$.

Now one can interpretate the “true” measurements $\hat{\mathbf{y}} \in \mathbb{R}^L$ as one realisation in $\hat{\omega}$ of y :

$$\hat{\mathbf{y}} = [y(x_1, \hat{\omega}), \dots, y(x_L, \hat{\omega})]^T. \quad (21)$$

Since uncertainties in measurements are inevitable, some measurement noise is added such that $\mathbf{z} := \hat{\mathbf{y}} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_L)^T$ is a centred Gaussian random vector with covariance $\mathbf{C}_{\boldsymbol{\epsilon}}$.

5 Discretisation

In order to numerically solve Eq. (12), one has to perform its full discretisation, in both the deterministic and stochastic spaces.

5.1 Spatial Discretisation

The spatial part of Eq. (15) is discretised by a standard finite element method. However, any other type of discretisation technique could be used. Since we deal with Galerkin methods in the stochastic space, assuming this also in the spatial domain gives a more compact representation of the problem. Taking a finite element ansatz $\mathcal{U}_N := \{\phi_n(x)\}_{n=1}^N \subset \mathcal{U}$ [47, 11, 56] as a corresponding subspace, the solution may be approximated by:

$$u(x, \omega) = \sum_{n=1}^N u_n(\omega) \phi_n(x), \quad (22)$$

where the coefficients $\{u_n(\omega)\}$ are now RVs in \mathcal{S} . Inserting the ansatz Eq. (22) back into Eq. (15) and applying the spatial Galerkin conditions [41, 38], one arrives at:

$$\mathbf{A}(\omega)[\mathbf{u}(\omega)] = \mathbf{f}(\omega), \quad (23)$$

where the parameter dependent symmetric and uniformly positive definite matrix $\mathbf{A}(\omega)$ is defined similarly to a usual finite element stiffness matrix as $(\mathbf{A}(\omega))_{m,n} := \mathbf{a}(\omega)(\phi_m, \phi_n)$ with the bi-linear form $\mathbf{a}(\omega)$ given by Eq. (16). Furthermore, the right hand side (r.h.s.) is determined by $(\mathbf{f}(\omega))_m := \langle \ell(\omega), \phi_m \rangle$ where the linear form $\ell(\omega)$ is given in Eq. (17), while $\mathbf{u}(\omega) = [u_1(\omega), \dots, u_N(\omega)]^T$ is introduced as a vector of random coefficients as in Eq. (22).

Eq. (23) represents a linear equation with random r.h.s. and random matrix. It is a semi-discretisation of some sort since it involves the variable ω and is still computationally intractable, as in general one needs infinitely many coordinates to parametrise Ω .

5.2 Stochastic Discretisation

The semi-discretised Eq. (23) is approximated such that the stochastic input data $\mathbf{A}(\omega)$ and $\mathbf{f}(\omega)$ are described with the help of RVs of some known type, here through a stochastic Galerkin (SG) method for the stochastic discretisation of Eq. (23) [15, 40, 24, 5, 53, 30, 41, 6, 53, 1, 54, 46]. Basic convergence of such an approximation may be established via Céa's lemma [41, 38].

In order to express the unknown coefficients (RVs) $u_n(\omega)$ in Eq. (22), we choose as ansatz functions multivariate *Hermite* polynomials $\{H_\alpha(\boldsymbol{\theta}(\omega))\}_{\alpha \in \mathcal{J}}$ in Gaussian RVs, also known under the name *Wiener's* polynomial chaos expansion (PCE) [28, 15, 40, 41, 38]

$$u_n(\boldsymbol{\theta}) = \sum_{\alpha \in \mathcal{J}} u_n^\alpha H_\alpha(\boldsymbol{\theta}(\omega)), \quad \text{or} \quad \mathbf{u}(\boldsymbol{\theta}) = \sum_{\alpha \in \mathcal{J}} \mathbf{u}^\alpha H_\alpha(\boldsymbol{\theta}(\omega)), \quad (24)$$

where $\mathbf{u}^\alpha := [u_1^\alpha, \dots, u_n^\alpha]^T$. The *Cameron-Martin* theorem assures us that the algebra of Gaussian variables is dense in $L_2(\Omega)$ [23, 34, 22, 20]. Here the index set \mathcal{J} is taken as a finite subset of $\mathbb{N}_0^{(\mathbb{N})}$, the set of all finite non-negative integer sequences, i.e. multi-indices. Although the set \mathcal{J} is finite with cardinality $|\mathcal{J}| = R$ and $\mathbb{N}_0^{(\mathbb{N})}$ is countable, there is no natural order on it; and hence we do not impose one at this point.

Inserting the ansatz Eq. (24) into Eq. (23) and applying the Bubnov-Galerkin projection onto the finite dimensional subspace $\mathcal{U}_N \otimes \mathcal{S}_{\mathcal{J}}$, one requires that the weighted residuals vanish:

$$\forall \beta \in \mathcal{J} : \quad \mathbb{E}([\mathbf{f}(\boldsymbol{\theta}) - \mathbf{A}(\boldsymbol{\theta})\mathbf{u}(\boldsymbol{\theta})]H_\beta(\boldsymbol{\theta})) = 0. \quad (25)$$

With $\mathbf{f}_\beta := \mathbb{E}(\mathbf{f}(\boldsymbol{\theta})H_\beta(\boldsymbol{\theta}))$ and $\mathbf{A}_{\beta,\alpha} := \mathbb{E}(H_\beta(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})H_\alpha(\boldsymbol{\theta}))$, Eq. (25) reads:

$$\forall \beta \in \mathcal{J} : \quad \sum_{\alpha \in \mathcal{J}} \mathbf{A}_{\beta,\alpha} \mathbf{u}^\alpha = \mathbf{f}_\beta, \quad (26)$$

which is a linear, symmetric and positive definite system of equations of size $N \times R$. The system is well-posed in the sense of Hadamard since the Lax-Milgram lemma applies on the subspace $\mathcal{U}_N \otimes \mathcal{S}_{\mathcal{J}}$.

To expose the structure of and compute the terms in Eq. (26), the parametric matrix in Eq. (23) is expanded in the Karhunen-Loève expansion (KLE) [41, 39, 16, 14] as

$$\mathbf{A}(\boldsymbol{\theta}) = \sum_{j=0}^{\infty} \mathbf{A}_j \xi_j(\boldsymbol{\theta}) \quad (27)$$

with scalar RVs ξ_j . Together with Eq. (15), it is not too hard to see that \mathbf{A}_j can be defined by the bilinear form

$$\mathbf{a}_j(v, u) := \int_{\mathcal{G}} \nabla v(x) \cdot (\kappa_j g_j(x) \nabla u(x)) \, dx, \quad (28)$$

with $g_j(x)$ being the coefficient of the KL expansion of $\kappa(x, \omega) = \sum_j \kappa_j g_j(x) \xi_j(\boldsymbol{\theta})$ and $(\mathbf{A}_j)_{m,n} := \mathbf{a}_j(\phi_m, \phi_n)$. These \mathbf{A}_j can be computed as “usual” finite element stiffness matrices with the “material properties” $\kappa_j g_j(x)$. It is worth noting that \mathbf{A}_0 is just the usual deterministic or mean stiffness matrix, obtained with the mean diffusion coefficient $\kappa_0(x)$ as parameter.

The parametric r.h.s. in Eq. (23) has an analogous expansion to Eq. (27), which may be either derived directly from the \mathbb{R}^N -valued RV $\mathbf{f}(\omega)$ — effectively a finite dimensional KLE — or from the continuous KLE of the random linear form in Eq. (17). In either case

$$\mathbf{f}(\omega) = \sum_i \varphi_i \psi_i(\omega) \mathbf{f}_i, \quad (29)$$

and, as in Eq. (27), only a finite number of terms are needed. For sparse representation of KLE see [25, 26]. The components in Eq. (26) may now be expressed as $\mathbf{f}_\beta = \sum_i \varphi_i f_\beta^i \mathbf{f}_i$ with $f_\beta^i := \mathbb{E}(H_\beta \psi_i)$. Observe that the random variables describing the input to the problem are $\{\xi_j\}$ and $\{\psi_i\}$.

Introducing the expansion Eq. (27) into Eq. (26) one obtains:

$$\forall \beta : \sum_{j=0}^{\infty} \sum_{\alpha \in \mathcal{J}} \Delta_{\beta, \alpha}^j \mathbf{A}_j \mathbf{u}^\alpha = \mathbf{f}_\beta, \quad (30)$$

where $\Delta_{\beta, \alpha}^j = \mathbb{E}(H_\beta \xi_j H_\alpha)$. Denoting the elements of the tensor product space $\mathbb{R}^N \otimes \mathbb{R}^R$ in an upright bold font as for example \mathbf{u} , and similarly linear operators on that space, as for example \mathbf{A} , one may further rewrite Eq. (30) in terms of tensor products [41, 38]:

$$\mathbf{A} \mathbf{u} := \left(\sum_{j=0}^{\infty} \mathbf{A}_j \otimes \Delta^j \right) \left(\sum_{\alpha \in \mathcal{J}} \mathbf{u}^\alpha \otimes \mathbf{e}^\alpha \right) = \left(\sum_{\alpha \in \mathcal{J}} \mathbf{f}_\alpha \otimes \mathbf{e}^\alpha \right) =: \mathbf{f}, \quad (31)$$

where \mathbf{e}^α denotes the canonical basis in \mathbb{R}^R . The tensor product is understood such that for $\mathbf{B} \in \mathbb{R}^{N \times N}$, $\mathbf{b} \in \mathbb{R}^N$, $\mathbf{G} \in \mathbb{R}^{R \times R}$, and $\mathbf{g} \in \mathbb{R}^R$, one has $(\mathbf{B} \otimes \mathbf{G})(\mathbf{b} \otimes \mathbf{g}) = (\mathbf{B} \mathbf{b}) \otimes (\mathbf{G} \mathbf{g})$. A concrete representation in terms of matrices and column vectors may be obtained by interpreting the symbol \otimes everywhere

as a *Kronecker* product. On the other hand, if \mathbf{u} is simply represented as the $N \times R$ matrix $\mathbf{u} = [\dots, \mathbf{u}^\alpha, \dots]$, then by exploiting the isomorphism between $\mathbb{R}^N \otimes \mathbb{R}^R$ and $\mathbb{R}^{N \times R}$ the term $(\mathbf{A}_j \otimes \mathbf{\Delta}^j)$ acts as $\mathbf{A}_j \mathbf{u} (\mathbf{\Delta}^j)^T$. With the help of Eq. (29) and the relations directly following it, the r.h.s. in Eq. (31) may be rewritten as

$$\mathbf{f} = \sum_{\alpha \in \mathcal{J}} \sum_i \varphi_i f_\alpha^i \mathbf{f}_i \otimes \mathbf{e}^\alpha = \sum_i \varphi_i \mathbf{f}_i \otimes \mathbf{g}^i, \quad (32)$$

where $\mathbf{g}^i := \sum_{\alpha \in \mathcal{J}} f_\alpha^i \mathbf{e}^\alpha$. Now the tensor product structure is exhibited also for the fully discrete counterpart to Eq. (15), and not only for the solution \mathbf{u} and r.h.s. \mathbf{f} , but also for the operator or matrix \mathbf{A} .

The operator \mathbf{A} in Eq. (31) inherits the properties of the operator in Eq. (15) in terms of symmetry and positive definiteness [41, 38]. The symmetry may be verified directly from Eq. (26), while the positive definiteness follows from the Galerkin projection and the uniform convergence in Eq. (31) on the finite dimensional space $\mathbb{R}^{(N \times N)} \otimes \mathbb{R}^{(R \times R)}$.

In order to make the procedure computationally feasible, of course the infinite sum in Eq. (27) has to be truncated at a finite value, say at M . The choice of M is now part of the stochastic discretisation and not an assumption. One simple guide for the choice of a minimal M is naturally the wish to introduce as much of the variance of the random field into the computation as possible. As the total variance σ_κ^2 is known, one may choose M such that the variance $\sigma_{\kappa, M}^2$ of the truncated field covers a desired fraction $\sigma_{\kappa, M}^2 / \sigma_\kappa^2$ of the total variance. Sometimes the truncated series may be scaled up by the inverse of the square root of that fraction to have again the full variance; in this way the first two moments of the random components describing the problem are correct, but of course higher moments are not.

Due to the uniform convergence alluded to above the sum can be extended far enough such that the operators \mathbf{A} in Eq. (31) are uniformly positive definite with respect to the discretisation parameters [41, 38]. This is in some way analogous to the use of numerical integration in the usual FEM [47, 11, 56].

The fully discrete forward problem may finally be announced as

$$\mathbf{A} \mathbf{u} = \left(\sum_{j=0}^M \mathbf{A}_j \otimes \mathbf{\Delta}^j \right) \mathbf{u} = \mathbf{f}. \quad (33)$$

Hence given $\kappa(x, \omega) := \exp(q(x, \omega))$ and $f(x, \omega)$, the matrix $\mathbf{A}(q)$ is the discrete form of the operator A in Eq. (1), and its inverse is the discrete form of the solution operator in Eq. (2).

6 Inverse Problem

On the system described by Eq. (1) we perform measurements formalised in Eq. (3). Typically these measurements z are only a shadow of the real quantity of interest q — like in Plato’s cave allegory — which we would like to identify, but are unable to observe. The measurements typically carry less information, and are disturbed by errors ϵ .

If we now view the parameters q due to the uncertainty as a random variable (with values in the space of admissible log-conductivity fields) and want to approximate this random variable with our previous knowledge — the prior random variable q_f (see Sec. 3) and the measurements z , then the minimum mean square estimator is one frequently used approximation. The estimator is also the minimum variance estimator, i.e. a linear Bayesian update with the variance of the difference as loss function. In the case of a linear problem and Gaussian random variables it is well known in the guise of the Kalman filter.

6.1 Linear Bayesian Estimator

We want to use Theorem 3.1, to update the information through projection, and we repeat the projected formula Eq. (10) in a matrix setting

$$\mathbf{q}_a(\omega) = \mathbf{q}_f(\omega) + \mathbf{K}(\mathbf{z}(\omega) - \mathbf{y}(\omega)), \quad (34)$$

with

$$\mathbf{K} = \mathbf{C}_{q_f y} (\mathbf{C}_y + \mathbf{C}_\epsilon)^{-1}. \quad (35)$$

By now — through the discretisation of the partial differential equation, the spaces \mathcal{Q} and \mathcal{Y} have become finite dimensional, so the Kalman gain \mathbf{K} in Eq. (35) is represented by a matrix. In the ensemble Kalman filter (EnKF) [12] this method is used in a way where \mathbf{q}_f , \mathbf{q}_a , ϵ , and \mathbf{y} are represented by Monte Carlo ensembles. Here, as already indicated at the end of Section 3, we want “to project the projection formula” Eq. (34) onto the polynomial chaos. The projection is simply computed — as the polynomial chaos is orthogonal — by multiplying Eq. (34) with each H_β , taking the expectation, and dividing by $\|H_\beta\|_{L_2(\Omega)}^2 = \beta!$, i.e. $\forall \beta$:

$$\mathbf{q}_a^\beta = \mathbb{E}(\mathbf{q}_a H_\beta) / \beta!$$

and so on for \mathbf{q}_f^β , \mathbf{z}^β , ϵ^β and \mathbf{y}^β . With this one has

$$\forall \beta : \quad \mathbf{q}_a^\beta = \mathbf{q}_f^\beta + \mathbf{K}(\mathbf{z}^\beta - \mathbf{y}^\beta), \quad (36)$$

where $\mathbf{z}^\beta = \hat{\mathbf{y}} \cdot \delta_{\beta,0} + \epsilon^\beta$, and $\delta_{\beta,\gamma}$ is the Kronecker symbol. This is again a tensorial equation, by setting for example as in Eq. (31):

$$\mathbf{q}_a = \sum_{\beta \in \mathcal{J}} \mathbf{q}_a^\beta \otimes \mathbf{e}^\beta, \quad (37)$$

and so on for \mathbf{z} and \mathbf{y} , where \mathbf{e}^β is the canonical basis in \mathbb{R}^R as before. The update Eq. (36) then may be written as

$$\mathbf{q}_a = \mathbf{q}_f + (\mathbf{K} \otimes \mathbf{I})(\mathbf{z} - \mathbf{y}). \quad (38)$$

Again the simplest representation of the tensor product is to collect all the column vectors into a matrix, e.g. $\mathbf{Q}_a = [..., \mathbf{q}_a^\beta, ...]$, $\mathbf{Z} = [..., \mathbf{z}^\beta, ...]$ and $\mathbf{Y} = [..., \mathbf{y}^\beta, ...]$ as indicated following Eq. (31). With this interpretation Eq. (38) reads

$$\mathbf{Q}_a = \mathbf{Q}_f + \mathbf{K}(\mathbf{Z} - \mathbf{Y}). \quad (39)$$

6.2 The Kalman filter as a special case

Let us remark that the term with $\beta = 0$ in Eq. (36), Eq. (38), or Eq. (39) is the update of the mean, thus recovering the well-known linear Bayes/ Kalman update for the mean. These equations also contain the Kalman update for the variance. For any expansion like

$$\mathbf{q}_a(\theta) = \sum_{\beta \in \mathcal{J}} \mathbf{q}_a^\beta H_\beta(\theta), \quad (40)$$

the variance is given by (see Appendix E)

$$\mathbf{C}_{q_a} = \mathbb{E}(\tilde{\mathbf{q}}_a(\cdot) \otimes \tilde{\mathbf{q}}_a(\cdot)) = \sum_{\gamma, \beta > 0} \mathbf{q}_a^\gamma \otimes \mathbf{q}_a^\beta \mathbb{E}(H_\gamma H_\beta) = \sum_{\gamma > 0} \mathbf{q}_a^\gamma \otimes \mathbf{q}_a^\gamma \gamma!, \quad (41)$$

as $\mathbb{E}(H_\gamma H_\beta) = \delta_{\gamma\beta} \gamma!$. Here $\tilde{\mathbf{q}}_a$ denotes \mathbf{q}_a with the first term equal to zero.

Defining the Gram matrix (in this case diagonal) $(\Delta^0)_{\gamma\beta} = \mathbb{E}(H_\gamma H_\beta) = \text{diag}(\gamma!)$, and using the matrix representation like in Eq. (39), Eq. (41) becomes

$$\mathbf{C}_{q_a} = \tilde{\mathbf{Q}}_a \Delta^0 \tilde{\mathbf{Q}}_a^T, \quad (42)$$

where $\tilde{\mathbf{Q}}_a$ is \mathbf{Q}_a with the $\gamma = 0$ term (the mean) missing. Using the assumption that $\mathbf{C}_{q_f \epsilon} = 0$, one may obtain from Eq. (39) the matrix equation:

$$\begin{aligned} \mathbf{C}_{q_a} &= \tilde{\mathbf{Q}}_a \Delta^0 \tilde{\mathbf{Q}}_a^T = \left((\tilde{\mathbf{Q}}_f + \mathbf{K}(\tilde{\mathbf{Z}} - \tilde{\mathbf{Y}})) \Delta^0 (\tilde{\mathbf{Q}}_f + \mathbf{K}(\tilde{\mathbf{Z}} - \tilde{\mathbf{Y}}))^T \right) \\ &= \mathbf{C}_{q_f} + \mathbf{K} \mathbf{C}_\epsilon \mathbf{K}^T + \mathbf{K} \mathbf{C}_y \mathbf{K}^T - \mathbf{C}_{q_f y} \mathbf{K}^T - \mathbf{K} \mathbf{C}_{q_f y}^T \\ &= \mathbf{C}_{q_f} + \mathbf{K}(\mathbf{C}_y + \mathbf{C}_\epsilon) \mathbf{K}^T - \mathbf{C}_{q_f y} \mathbf{K}^T - \mathbf{K} \mathbf{C}_{q_f y}^T \end{aligned} \quad (43)$$

With \mathbf{K} from Eq. (35), one obtains

$$\begin{aligned} \mathbf{C}_{q_a} &= \mathbf{C}_{q_f} + \mathbf{C}_{q_f y} (\mathbf{C}_y + \mathbf{C}_\epsilon)^{-1} \mathbf{C}_{q_f y}^T - 2\mathbf{C}_{q_f y} (\mathbf{C}_y + \mathbf{C}_\epsilon)^{-1} \mathbf{C}_{q_f y}^T \\ &= \mathbf{C}_{q_f} - \mathbf{C}_{q_f y} (\mathbf{C}_y + \mathbf{C}_\epsilon)^{-1} \mathbf{C}_{q_f y}^T \end{aligned} \quad (44)$$

This is exactly the usual Kalman filter update for the variance estimate.

Notice that in Eq. (36), Eq. (38), or Eq. (39) the complete random variable is updated — up to the order kept in the PCE — and not just the first two moments, as in the usual case of the Gauss-Markov theorem (for example in the guise of the Kalman filter).

7 Algorithm

The linear Bayesian update procedure described in this paper can be implemented as is shown in Algorithm 1.

Linear Bayesian Update

Input: a priori $\mathbf{q}_f(\omega)$ and measurement $\mathbf{z}(\omega)$

- approximate a priori information $\mathbf{q}_f(\omega)$ and input $\mathbf{z}(\omega)$ by PCE:

$$\mathbf{Q}_f = [\dots, \mathbf{q}_f^\beta, \dots], \mathbf{Z} = [\dots, \mathbf{z}^\beta, \dots];$$

- centralise \mathbf{Z} to $\tilde{\mathbf{Z}}$ (take out the mean);
- solve stochastic forward problem

$$\mathbf{u}(\omega) = S(\mathbf{q}_f(\omega); \mathbf{f}(\omega));$$

- forecast of measurement

$$\mathbf{y}(\omega) = Y(\mathbf{q}_f(\omega); \mathbf{u}(\omega)) = Y(\mathbf{q}_f(\omega); S(\mathbf{q}_f(\omega); \mathbf{f}(\omega)));$$

- PCE representation of $\mathbf{y}(\omega)$:

$$\mathbf{Y} = [\dots, \mathbf{y}^\beta, \dots];$$

- centralise \mathbf{Y} to $\tilde{\mathbf{Y}}$;
- compute covariance (see Eq. (41))

$$\mathbf{C}_d = \mathbf{C}_y + \mathbf{C}_\epsilon = \tilde{\mathbf{Y}} \mathbf{\Delta}^0 \tilde{\mathbf{Y}}^T + \mathbf{C}_\epsilon;$$

- solve with an appropriate method for \mathbf{G} (e.g., QR or SVD)

$$\mathbf{G} = \mathbf{C}_d^{-1} (\mathbf{Z} - \mathbf{Y});$$

- compute covariance

$$\mathbf{C}_{q_f y} = \tilde{\mathbf{Q}}_f \mathbf{\Delta}^0 \tilde{\mathbf{Y}}^T;$$

- compute formula in Eq. (39)

$$\mathbf{Q}_a = \mathbf{Q}_f + \mathbf{C}_{q_f y} \mathbf{G};$$

Output: assimilated data $\mathbf{Q}_a = [\dots, \mathbf{q}_a^\beta, \dots]$

Algorithm 1: Implementation of a general linear Bayesian update based on PCE

The inputs are defined as a priori information $\mathbf{q}_f(\omega)$ and the given measurements \mathbf{z} , which are assumed to be obtained by experiments. In this paper the measurements are simulated, and modelled as uncertain due to existing measurement noise ϵ . In order to identify the posterior with the help of previously described direct update procedure, the input data are transformed to a corresponding polynomial chaos expansions. These data are then stored in a matrix format \mathbf{Q}_f and \mathbf{Z} respectively, where each column contains the vector of a certain PCE coefficients. For the sake of simplicity the measurement error is assumed centred Gaussian with covariance $\mathbf{C}_\epsilon := \sigma_\epsilon^2 \mathbf{I}$. With the PCE of the a priori parameter one may solve the forward problem and obtain the PCE of the solution \mathbf{u} , which allows the calculation of some functional of the solution — the forecast of the measurement — \mathbf{y} . Subtracting the mean value, one may compute the covariance \mathbf{C}_y (see Eq. (41)), then find the pseudo-inverse of \mathbf{C}_d and multiply it by the difference $(\mathbf{Z} - \mathbf{Y})$. Further, it is easy to find the covariance $\mathbf{C}_{q_f y}$ between the a priori and simulated data and to compute the formula Eq. (39) in order to get the matrix of PCE coefficients of the posterior \mathbf{Q}_a (see Appendix E).

Due to the absence of real experiments the measurement data in this paper are simulated as shown in Algorithm 2. Namely, taking some deterministic values for the parameters and deterministic loading conditions, we obtain the solution \mathbf{u} (i.e. \mathbf{y}). As such obtained values are 'exact', we disturb them by a previously described noise ϵ , to simulate a “real” measurement.

Virtual Reality — Simulation of a Measurement

Input: $q(x), f(x), x \in \mathcal{G}$

- Conductivity $\kappa(x) = \exp(q(x))$;
- solve for $u(x)$ from:
 - $\text{div}(\kappa(x)\nabla u(x)) = f(x), \forall x \in \mathcal{G}$;
 - $u(x) = 0, \forall x \in \partial\mathcal{G}$;
- Compute “true” measurements
 - $\hat{y}_i = \int_{\mathcal{G}_i} u(x)dx, \quad \forall \mathcal{G}_i \in \hat{\mathcal{G}}, \quad i = \{1, 2, \dots, L\}$;
- Generate measurement noise
 - here $\epsilon \sim N(0, \mathbf{C}_\epsilon) \in \mathbb{R}^L$;
- Compute noisy measurements
 - $\mathbf{z} = \hat{\mathbf{y}} + \epsilon$;
- Approximate \mathbf{z} by truncated PCE
 - $\mathbf{z} = \sum_{\beta \in \mathcal{J}} \mathbf{z}^\beta H_\beta$;

Output: $\mathbf{Z} = [\dots, \mathbf{z}^\beta, \dots]$

Algorithm 2: Simulation of measurement

The step of solving the forward problem may be performed in several ways. Here the so-called low-rank sparse stochastic Galerkin method [42, 55] and tensorial algebra as it is provided in [27] were used.

8 Numerical Results

The direct Bayesian update method is tested on the example introduced in Section 4, defined on a bounded L -shaped domain. The numerical results are analysed with respect to the known analytic solution, the “truth” used in Algorithm 2.

8.1 The Measurement and the “True” Value of κ

Reality — the “truth” — (see Algorithm 2) is taken in three different scenarios, where the “true” conductivity parameter is assumed as either:

1. a constant value over the spatial domain \mathcal{G} , here $\kappa = 2$;
2. a linear function of the position $\mathbf{x} \in \mathcal{G}$, here $\kappa = 2 + 0.3 \cdot (x + y)$;
3. a quadratic function in terms of $\mathbf{x} \in \mathcal{G}$, here $\kappa = 2.2 - 0.1 \cdot (x^2 + y^2)$.

Note that these models are not intended to describe a specific real system but are rather set up to demonstrate the method. Furthermore, we choose the right hand side in all experiments as a deterministic sinusoidal function $f = f_0 \sin(\frac{2\pi}{\lambda} \mathbf{x}^T \mathbf{d} + \varphi)$ where f_0 represents the amplitude, λ the wave-length, φ the phase, and $\mathbf{d} = [\cos \alpha \sin \alpha]$ the direction of the sinusoidal wave specified by an angle $\alpha \in [-\pi/2, \pi/2]$. The values of these parameters are specific to each experiment and are given later in the text.

The forward problem is solved within the finite element framework by discretising the spatial domain into 1032 triangular elements. In this way, the hydraulic head data $\hat{\mathbf{y}}$ are obtained on the measurement points, chosen to be uniformly distributed over the whole domain except the boundary nodes, which have been excluded from the measurement as a trivial case.

8.2 Stochastic Forward Problem Example

The stationary diffusion equation may be cast as a forward model that predicts the value of the hydraulic head field at specific locations from the conductivity field. The initial assumption — the a priori distribution — on this parameter is given by a lognormal random field described by a spatially constant mean value $\bar{\kappa} = 2.4$ and standard deviation $\sigma_\kappa = 0.4$. The field

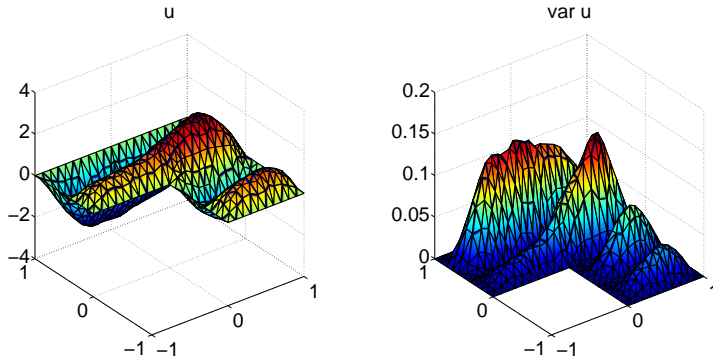


Figure 1: Mean value \bar{u} and variance of simulated data u

is a nonlinear transformation of a Gaussian random field described by an exponential covariance function and correlation length $l_c = 1$ as it is given in Eq. (13). It is approximated through a KL-expansion in 50 terms, followed by a polynomial chaos expansion of order 3 [41]. The spatial discretisation is the same as in the simulation of the measurement.

As an example of the stochastic forward response for the a priori distribution the mean value and standard deviation of the solution for the right hand side $f = \sin(2\pi x + \pi/8)$ are displayed in Fig. 1.

8.2.1 Update Procedure

The experiment is set up by averaging the hydraulic head u over a number of patches, the centres of which are shown in Tab. 1. This constitutes the “true” measurement \hat{y} .

The measurements are often performed several times by applying each time different loading conditions which come from a change of the wave-length λ , the phase $\varphi = [0, 2\pi]$, and the direction \mathbf{d} (by changing $\alpha \in [-\pi/2, \pi/2]$) of the sinusoidal wave. Due to this the parameter estimation is repeated here in a sequential process (see Tab. 2). Each series of updates is an experiment, which is performed independently for a different number of measurement points, to show the influence of the amount of the measured information. Thus in each experiment the posterior from the first update represents the prior for the second update and so on.

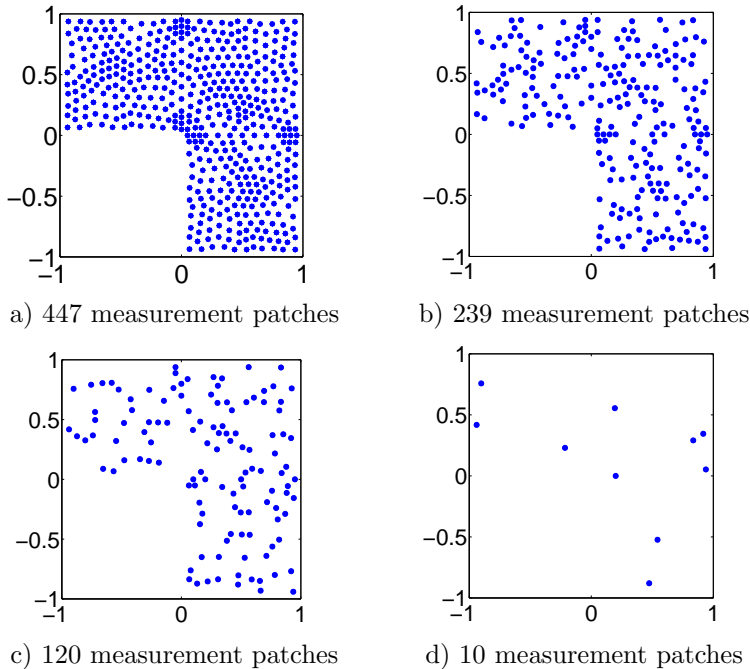


Table 1: Position of measurement points (FEM nodes) used in the experiments

Exp.	L	ε_p	ε_a : 1st up.	2nd up.	3rd up.	4th up.
1.	477	0.45	0.08	0.04	0.03	0.03
2.	239	0.45	0.08	0.05	0.05	0.04
3.	120	0.45	0.07	0.05	0.05	0.04
4.	60	0.45	0.07	0.06	0.05	0.05
5.	10	0.45	0.13	0.08	0.07	0.07

Table 2: “**Constant truth**”: Decay of the relative error ε_a in each experiment

In order to describe the properties of the update procedure, let us define the relative errors ε_a and $\bar{\varepsilon}_a$ corresponding to the posterior distribution after each update:

$$\varepsilon_a := \frac{\|\kappa_a - \kappa_t\|_{L_2(\Omega \otimes \mathcal{G})}}{\|\kappa_t\|_{L_2(\Omega \otimes \mathcal{G})}}; \quad \bar{\varepsilon}_a := \frac{|\mathbb{E}(\kappa_a) - \mathbb{E}(\kappa_t)|}{|\mathbb{E}(\kappa_t)|} \quad (45)$$

where κ_t represents the truth. The corresponding errors for the initial prior are denoted by ε_p and $\bar{\varepsilon}_p$ respectively. With this we define errors ratios $\rho := \varepsilon_a/\varepsilon_p$ and $\bar{\rho} := \bar{\varepsilon}_a/\bar{\varepsilon}_p$: which define the quantity:

$$I = 1 - \rho, \quad (46)$$

here called the total *improvement* of the update. This quantity describes the relation between the prior and posterior, i.e., how much the posterior has improved after the update procedure from the prior compared to the “truth”.

One may follow the change of ε_a of the update in Tab. 2 with respect to the number of experiments as well as with respect to the number of measurement points L . As expected, the error is smaller in experiment with a higher number of measurement points. It also decreases in each experiment with the number of sequential updates. This happens due to additional available information in the update process. However, it sometimes may happen that the error with a smaller set of measurement points is better than with the larger one. The reason are the measurement point positions, which in some

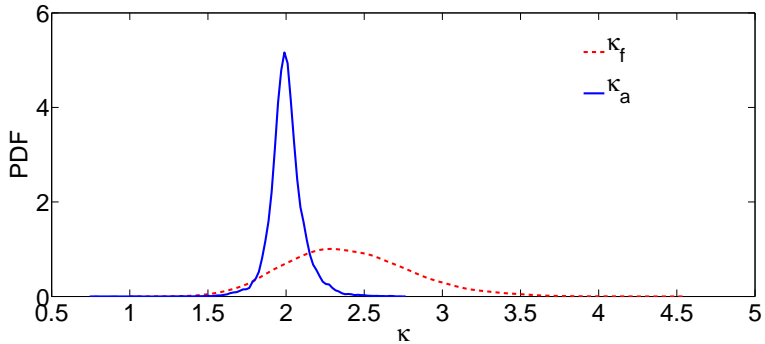


Figure 2: “Constant truth”, experiment 3 ($L=120$): Posterior probability density function κ_a compared to the prior κ_f for a single point in domain

region of the domain bring much more information into the update process

than in others. For example the boundary nodes bring trivial information into the update process and hence we exclude them. In the case of the “constant truth” already ca. 10 measurement points suffice after four updates to reduce the error to 7%. We have not tried to optimise the position of the measurement points/patches or the loading in the experiments, which is beyond the scope of this paper.

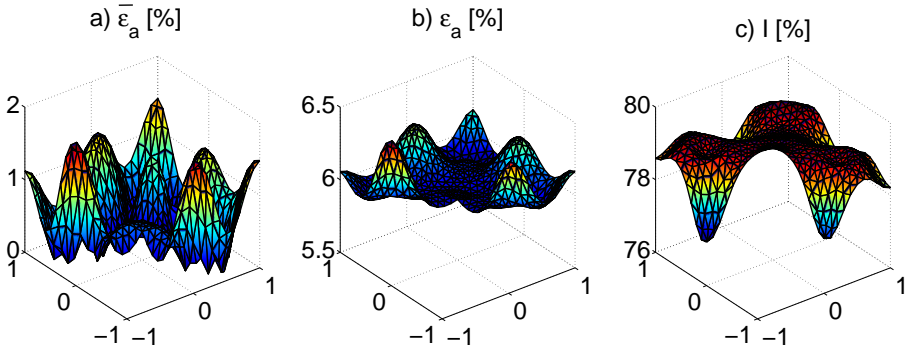


Figure 3: “Constant truth”, experiment 1 (L=447) after 4th update: a) Relative error $\bar{\varepsilon}_a$ (the mean of the posterior compared to the mean of the truth) b) relative error ε_a (the posterior compared to the truth) c) improvement I (the posterior compared to the prior)

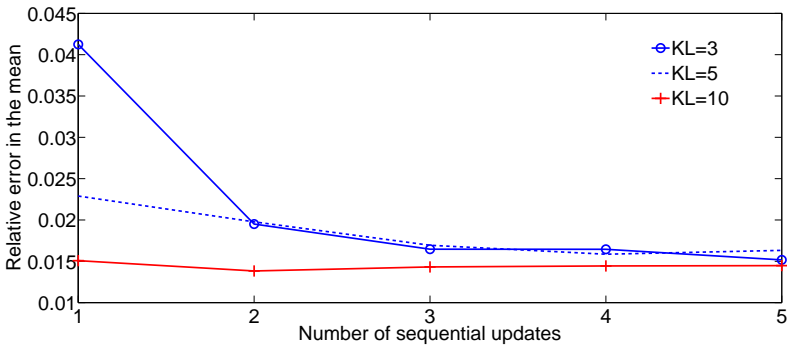


Figure 4: “Constant truth”, experiment 1 (L=447): Convergence of the relative error in the mean of the posterior, $\bar{\varepsilon}_a$, with the number of sequential updates and a different number of KL terms in the expansion of the prior

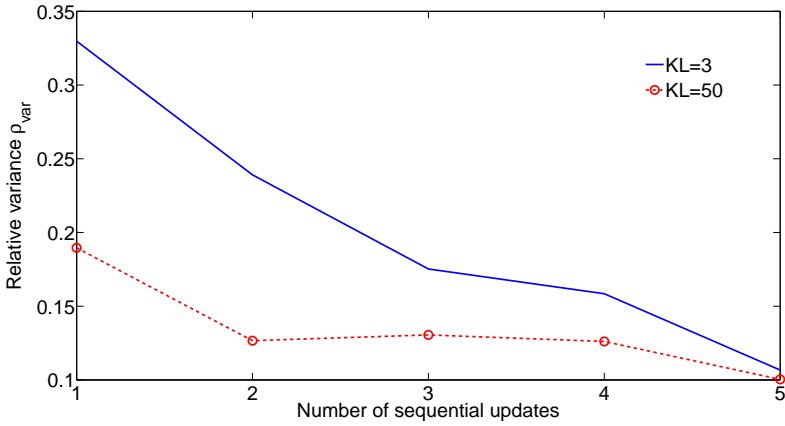


Figure 5: **“Constant truth”, experiment 1 (L=447)**: Convergence of ρ_{var} with the number of updates and the measurement points

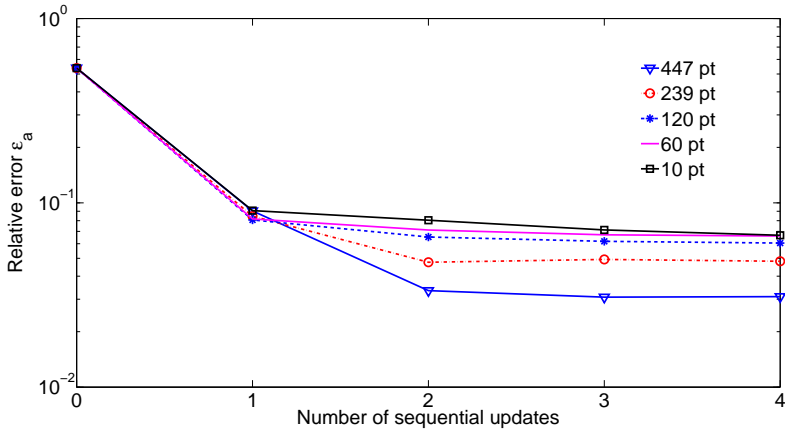


Figure 6: **“Linear truth”, experiment 1 (L=447)**: Convergence behaviour of the relative error ε_a with respect to the number of sequential updates and measurement points

Furthermore, comparing the probability density function of the a priori and a posteriori distribution in one point of the domain (experiment 3.) in Fig. 2, one may notice that the latter one is more narrowed and centred around the “true” value 2. This is also shown in Fig. 3a, where the relative error $\bar{\varepsilon}_a$ in the mean already after the first update reduces to 2%, which means that the first order moment is almost instantaneously identified.

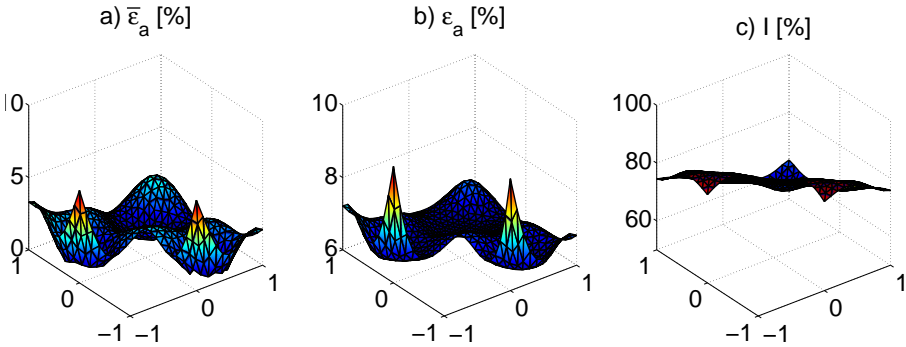


Figure 7: “**Linear truth**”, experiment 1 (L=447) after 4th update: a) Relative error $\bar{\varepsilon}_a$ (the mean of the posterior compared to the truth) b) relative RMS error ε_a (the posterior compared to the truth) c) improvement I (the posterior compared to the prior)

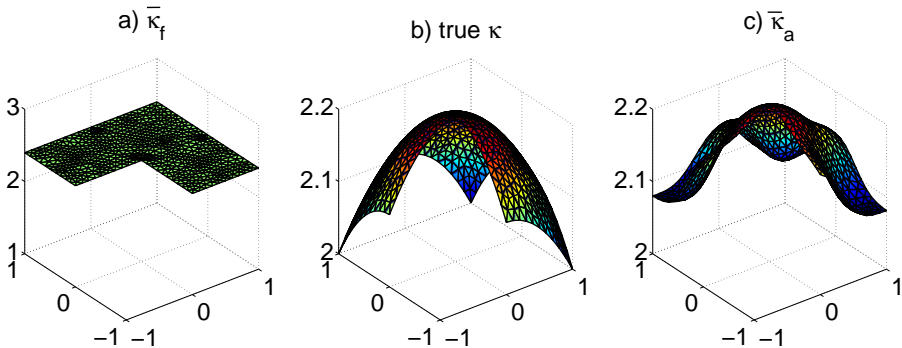


Figure 8: “**Quadratic truth**”, experiment 1 (L=447) after 4th update: a) mean of the prior, $\bar{\kappa}_f$ b) truth, κ c) mean of the posterior, $\bar{\kappa}_a$

However, the error in higher order terms (see Fig. 3b) is more important

since it describes how much of the uncertainty has been removed. This is also shown in Fig. 3c, where we plot the improvement I from Eq. (46) obtained after the first update in the first experiment. In addition, we study the behaviour of the mean and variance of the posterior with respect to the number of KLE modes used in the expansion of the prior. The mean value converges relatively fast (see Fig. 4). Already with 10 KLE modes we obtain an error of 1%. However, the variance converges a bit slower. One needs 50 KLE modes in order to significantly reduce the error in variance. This we show by plotting the convergence of the relative variance ρ_{var} in Fig. 5 obtained as a ratio between the posterior and prior variance.

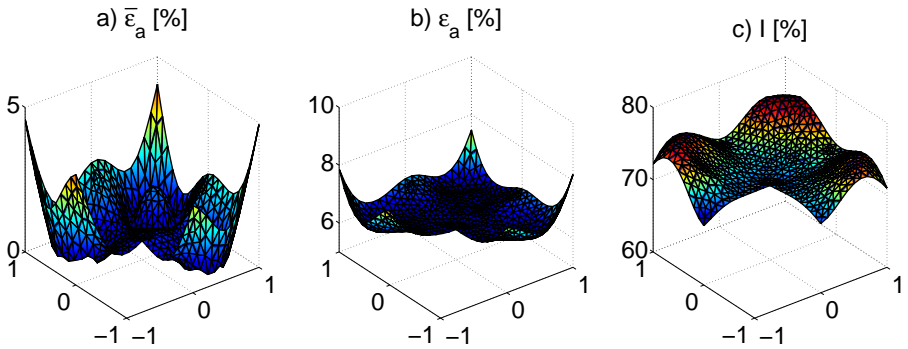


Figure 9: “Quadratic truth”, experiment 1 ($L=447$) after 4th update
a) Relative error $\bar{\varepsilon}_a$ (the mean of the posterior compared to the truth) b)
relative RMS error ε_a (the posterior compared to the truth) c) improvement
 I (the posterior compared to the prior)

For the “linear truth”, we repeat the same procedure as given in Tab. 2, i.e. five independent experiments, each of them containing four sequential updates. The convergence of the relative error has similar properties as before as one may see in Fig. 6. Thus, the error ratio ρ drops with the number of updates, as it is expected, due to increasing precision of the a priori distribution compared to the truth.

The behaviour of the relative error ε_a is similar to the case of the “constant truth”, see Fig. 7. However, a problem arises in the identification of the boundary values, where the relative error and the error in the mean become larger and hence the improvement smaller. This is an expected behaviour since the boundary nodes are excluded from the measurement as a trivial case (zero boundary conditions).

In Fig. 8, we reconstruct a “quadratic truth” from a spatially uniform

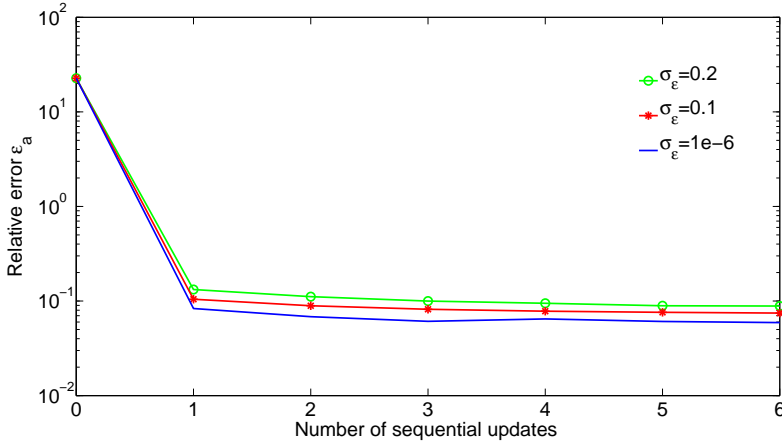


Figure 10: **“Quadratic truth”, experiment 1 (L=447)**: Convergence of the relative error with the number of updates and the measurement error σ_ϵ obtained for 447 measurement points

lognormal prior. The error behaviour in Fig. 9 is similar to the previous case of the “linear truth”.

Having in mind that the measurements are disturbed with centred Gaussian noise with standard deviation σ_ϵ , we have investigated the influence of this parameter on the update process in Fig. 10. Here one may learn that this parameter has only a minor influence on the update. Obviously the main contribution to the error for the range of tested σ_ϵ still comes from the incomplete information in the measurement.

9 Conclusion

A linear Bayesian estimation of unknown parameters is formulated in a purely deterministic and algebraic way, without the need for any kind of sampling techniques like MCMC. The regularisation of the ill-posed problem is achieved by introduction of a priori information approximated by a combination of Karhunen-Loève and polynomial chaos expansions truncated to a finite number of terms. This representation enters a stochastic forward model, solved by a Galerkin procedure in a low-rank and sparse format. Taking the forward solution to forecast the measurement, the update of the prior is a projection of the minimum variance estimator from linear Bayesian updating

onto the polynomial chaos basis. The new estimate is obtained for all terms of the PCE, i.e. one is able to estimate all higher order moments, not just the mean and the variance. We have shown that for the mean and variance the estimation is of the Kalman type.

The suggested method is tested on the stationary diffusion equation with prescribed source terms and unknown conductivity parameter. The identification experiment, characterised by a sequential update process, is performed for a “deterministic truth” with increasing spatial compability. In this way we have shown the influence of additional information and loading conditions on the update process.

The presented linear Bayesian update does not need any linearity in the forward model, and it can readily update non-Gaussian uncertainties. In addition, it accommodates noisy measurements and skewed RVs.

Acknowledgement: The support of the Deutsche Forschungsgemeinschaft (DFG), German Luftfahrtforschungsprogramm, the Federal Ministry of Economics (BMWA), and SPT Group GmbH in Hamburg is gratefully acknowledged.

Appendix

These appendices list some fundamental properties of the polynomial chaos expansion (PCE), the connected Hermite algebra, and the use of the Hermite transform.

A Multi-Indices

In the above formulation, the need for multi-indices of arbitrary length arises. Formally they may be defined by

$$\alpha = (\alpha_1, \dots, \alpha_j, \dots) \in \mathbb{N}_0^{(\mathbb{N})} := \mathcal{N}, \quad (47)$$

which are sequences of non-negative integers, only finitely many of which are non-zero. As by definition $0! := 1$, the following expressions are well defined:

$$|\alpha| := \sum_{j=1}^{\infty} \alpha_j, \quad \alpha! := \prod_{j=1}^{\infty} \alpha_j!, \quad \ell(\alpha) := \max\{j \in \mathbb{N} \mid \alpha_j > 0\}. \quad (48)$$

B Hermite Polynomials

As there are different ways to define — and to normalise — the Hermite polynomials, a specific way has to be chosen. In applications with probability theory it seems most advantageous to use the following definition [20, 22, 23, 34]:

$$h_k(t) := (-1)^k e^{t^2/2} \left(\frac{d}{dt} \right)^k e^{-t^2/2}; \quad \forall t \in \mathbb{R}, k \in \mathbb{N}_0, \quad (49)$$

where the coefficient of the highest power of t — which is t^k for h_k — is equal to unity.

The first five polynomials are:

$$\begin{aligned} h_0(t) &= 1, & h_1(t) &= t, & h_2(t) &= t^2 - 1, \\ h_3(t) &= t^3 - 3t, & h_4(t) &= t^4 - 6t^2 + 3. \end{aligned}$$

The recursion relation for these polynomials is

$$h_{k+1}(t) = t h_k(t) - k h_{k-1}(t); \quad k \in \mathbb{N}. \quad (50)$$

These are orthogonal polynomials w.r.t standard Gaussian probability measure Γ , where $\Gamma(dt) = (2\pi)^{-1/2} e^{-t^2/2} dt$ — the set $\{h_k(t)/\sqrt{k!} \mid k \in \mathbb{N}_0\}$

forms a complete orthonormal system (CONS) in $L_2(\mathbb{R}, \Gamma)$ — as the Hermite polynomials satisfy

$$\int_{-\infty}^{\infty} h_m(t) h_n(t) \Gamma(dt) = n! \delta_{nm}. \quad (51)$$

Multi-variate Hermite polynomials will be defined right away for an infinite number of variables, i.e. for $\mathbf{t} = (t_1, t_2, \dots, t_j, \dots) \in \mathbb{R}^{\mathbb{N}}$, the space of all sequences. This uses the multi-indices defined in Appendix A: For $\alpha = (\alpha_1, \dots, \alpha_j, \dots) \in \mathcal{N}$ remember that except for a finite number all other α_j are zero; hence in the definition of the multi-variate Hermite polynomial

$$H_{\alpha}(\mathbf{t}) := \prod_{j=1}^{\infty} h_{\alpha_j}(t_j); \quad \forall \mathbf{t} \in \mathbb{R}^{\mathbb{N}}, \alpha \in \mathcal{N}, \quad (52)$$

except for finitely many factors all others are h_0 , which equals unity, and the infinite product is really a finite one and well defined.

The space $\mathbb{R}^{\mathbb{N}}$ can be equipped with a Gaussian (product) measure [20, 22, 23, 34], again denoted by Γ . Then the set $\{H_{\alpha}(\mathbf{t})/\sqrt{\alpha!} \mid \alpha \in \mathcal{N}\}$ is a CONS in $L_2(\mathbb{R}^{\mathbb{N}}, \Gamma)$ as the multivariate Hermite polynomials satisfy

$$\int_{\mathbb{R}^{\mathbb{N}}} H_{\alpha}(\mathbf{t}) H_{\beta}(\mathbf{t}) \Gamma(d\mathbf{t}) = \alpha! \delta_{\alpha\beta}, \quad (53)$$

where the Kronecker symbol is extended to $\delta_{\alpha\beta} = 1$ in case $\alpha = \beta$ and zero otherwise.

C The Hermite Algebra

Consider first the usual univariate Hermite polynomials $\{h_k\}$ as defined in Appendix B, Eq. (49). As the univariate Hermite polynomials are a linear basis for the polynomial algebra, i.e. every polynomial can be written as linear combination of Hermite polynomials, this is also the case for the product of two Hermite polynomials $h_k h_{\ell}$, which is clearly also a polynomial:

$$h_k(t) h_{\ell}(t) = \sum_{n=|k-\ell|}^{k+\ell} c_{k\ell}^{(n)} h_n(t) \quad (54)$$

The coefficients are only non-zero [34] for integer $g = (k + \ell + n)/2 \in \mathbb{N}$ and if $g \geq k \wedge g \geq \ell \wedge g \geq n$. They can be explicitly given

$$c_{k\ell}^{(n)} = \frac{k! \ell!}{(g-k)! (g-\ell)! (g-n)!}, \quad (55)$$

and are called the structure constants of the univariate Hermite algebra.

For the multivariate Hermite algebra, analogous statements hold [34]:

$$H_\alpha(\mathbf{t})H_\beta(\mathbf{t}) = \sum_{\gamma} c_{\alpha\beta}^{\gamma} H_{\gamma}(\mathbf{t}). \quad (56)$$

with the multivariate structure constants

$$c_{\alpha\beta}^{\gamma} = \prod_{j=1}^{\infty} c_{\alpha_j\beta_j}^{\gamma_j}, \quad (57)$$

defined in terms of the univariate structure constants Eq. (55).

From this it is easy to see that

$$\mathbb{E}(H_\alpha H_\beta H_\gamma) = \mathbb{E}\left(H_\gamma \sum_{\epsilon} c_{\alpha\beta}^{\epsilon} H_{\epsilon}\right) = c_{\alpha\beta}^{\gamma} \gamma!. \quad (58)$$

Products of more than two Hermite polynomials may be computed recursively, we here look at triple products as an example, using Eq. (56):

$$H_\alpha H_\beta H_\delta = \left(\sum_{\gamma} c_{\alpha\beta}^{\gamma} H_{\gamma}\right) H_{\delta} = \sum_{\epsilon} \left(\sum_{\gamma} c_{\gamma\delta}^{\epsilon} c_{\alpha\beta}^{\gamma}\right) H_{\epsilon}. \quad (59)$$

D The Hermite Transform

A variant of the Hermite transform maps a random variable onto the set of expansion coefficients of the PCE [22]. Any random variable $r \in L_2(\Omega, \mathcal{V})$ which may be represented with a PCE

$$r(\omega) = \sum_{\alpha \in \mathbb{N}_0^{(N)}} \varrho^{\alpha} H_{\alpha}(\theta(\omega)), \quad (60)$$

is mapped onto

$$\mathcal{H}(r) := (\varrho^{\alpha})_{\alpha \in \mathcal{N}} = (\varrho) \in \mathbb{R}^{\mathcal{N}}. \quad (61)$$

This way $\bar{r} := \mathbb{E}(r) = \varrho^0$ and $\mathcal{H}(\bar{r}) = (\varrho^0, 0, 0, \dots)$, as well a $\tilde{r}(\omega) = r(\omega) - \bar{r}$ and $\mathcal{H}(\tilde{r}) = (0, (\varrho^{\alpha})_{\alpha \in \mathcal{J}, \alpha > 0})$ are explicitly given. These sequences may be seen also as the coefficients of power series in infinitely many complex variables $\mathbf{z} \in \mathbb{C}^{\mathbb{N}}$, namely by

$$\sum_{\alpha \in \mathbb{N}_0^{(N)}} \varrho^{\alpha} \mathbf{z}^{\alpha},$$

where $\mathbf{z}^\alpha := \prod_j z_j^{\alpha_j}$. This is the original definition of the Hermite transform [22].

It can be used to easily compute the Hermite transform of the ordinary product like in Eq. (56), as

$$\mathcal{H}(H_\alpha H_\beta) = (c_{\alpha\beta}^\gamma)_{\gamma \in \mathbb{N}_0^{(N)}}. \quad (62)$$

With the structure constants Eq. (57) one defines the matrices $\mathbf{Q}_2^\gamma := (c_{\alpha\beta}^\gamma)$ with indices α and β . With this notation the Hermite transform of the product of two random variables $r_1(\omega) = \sum_{\alpha \in \mathbb{N}_0^{(N)}} \varrho_1^\alpha H_\alpha(\theta)$ and $r_2(\omega) = \sum_{\beta \in \mathbb{N}_0^{(N)}} \varrho_2^\beta H_\beta(\theta)$ is

$$\mathcal{H}(r_1 r_2) = ((\varrho_1) \mathbf{Q}_2^\gamma (\varrho_2)^T)_{\gamma \in \mathbb{N}_0^{(N)}} \quad (63)$$

Each coefficient is a bilinear form in the coefficient sequences of the factors, and the collection of all those bilinear forms $\mathbf{Q}_2 = (\mathbf{Q}_2^\gamma)_{\gamma \in \mathbb{N}_0^{(N)}}$ is a bilinear mapping that maps the coefficient sequences of r_1 and r_2 into the coefficient sequence of the product

$$\mathcal{H}(r_1 r_2) =: \mathbf{Q}_2((\varrho_1), (\varrho_2)) = \mathbf{Q}_2(\mathcal{H}(r_1), \mathcal{H}(r_2)). \quad (64)$$

Products of more than two random variables may now be defined recursively through the use of associativity. e.g. $r_1 r_2 r_3 r_4 = (((r_1 r_2) r_3) r_4)$:

$$\begin{aligned} \forall k > 2: \quad \mathcal{H} \left(\prod_{j=1}^k r_j \right) &:= \mathbf{Q}_k((\varrho_1), (\varrho_2), \dots, (\varrho_k)) := \\ &\mathbf{Q}_{k-1}(\mathbf{Q}_2((\varrho_1), (\varrho_2)), (\varrho_3) \dots, (\varrho_k)). \end{aligned} \quad (65)$$

Each \mathbf{Q}_k is again composed of a sequence of k -linear forms $\{\mathbf{Q}_k^\gamma\}_{\gamma \in \mathbb{N}_0^{(N)}}$, which define each coefficient of the Hermite transform of the k -fold product.

E Higher order moments

Consider RVs $\mathbf{r}_j(\omega) = \sum_{\alpha \in \mathcal{N}} \rho_j^\alpha H_\alpha(\theta)$ with values in a vector space \mathcal{V} (see Appendix D), then $\bar{\mathbf{r}}$, $\tilde{\mathbf{r}}(\omega)$, as well as ρ_j^α are in \mathcal{V} . More generally, any moment may be easily computed knowing the PCE. The k -th centred moment of the RVs $\mathbf{r}_1(\omega), \mathbf{r}_2(\omega), \dots, \mathbf{r}_k(\omega)$ is defined as

$$\mathbf{M}_{\mathbf{r}_1 \dots \mathbf{r}_k}^k = \mathbb{E} \left(\otimes_{j=1}^k \tilde{\mathbf{r}}_j \right), \quad (66)$$

a tensor of order k . Hence the k -point correlation (the k -th moment) may be expressed as

$$\mathbf{M}_{\mathbf{r}_1 \dots \mathbf{r}_k}^k = \sum_{\gamma^1, \dots, \gamma^k > 0} \mathbb{E} \left(\prod_{j=1}^k H_{\gamma^j}(\boldsymbol{\theta}) \right) \bigotimes_{m=1}^k \boldsymbol{\rho}_m^{\gamma^j}, \quad (67)$$

and in particular:

$$\mathbf{C}_{\mathbf{r}_1 \mathbf{r}_2} = \mathbf{M}_{\mathbf{r}_1 \mathbf{r}_2}^2 = \mathbb{E}(\tilde{\mathbf{r}}_1 \otimes \tilde{\mathbf{r}}_2) = \sum_{\gamma, \beta > 0} \boldsymbol{\rho}_1^\gamma \otimes \boldsymbol{\rho}_2^\beta \mathbb{E}(H_\gamma H_\beta) = \sum_{\gamma > 0} \boldsymbol{\rho}_1^\gamma \otimes \boldsymbol{\rho}_2^{\gamma!}, \quad (68)$$

as $\mathbb{E}(H_\gamma H_\beta) = \delta_{\gamma\beta} \gamma!$. The expected values of the products of Hermite polynomials in Eq. (67) may be computed analytically, using the formulas from Appendix C and D.

References

- [1] S. Acharjee and N. Zabaras. A non-intrusive stochastic Galerkin approach for modeling uncertainty propagation in deformation processes. *Computers & Structures*, 85:244–254, 2007.
- [2] M. Arnst, R. Ghanem, and C. Soize. Identification of Bayesian posteriors for coefficients of chaos expansions. *Journal of Computational Physics*, 229(9):3134 – 3154, 2010.
- [3] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2006.
- [4] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421, 2006.
- [5] I. Babuška, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, 2004.
- [6] I. Babuška, R. Tempone, and G. E. Zouraris. Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation. *Computer Methods in Applied Mechanics and Engineering*, 194(12-16):1251–1294, 2005.

- [7] S. Balakrishnan, A. Roy, M. G. Ierapetritou, G. P. Flach, and P. G. Georgopoulos. Uncertainty reduction and characterization for complex environmental fate and transport models: An empirical Bayesian framework incorporating the stochastic response surface method. *Water Resources Research*, 39(12):1350–1362, December 2003.
- [8] E. D. Blanchard. *Polynomial Chaos Approaches to Parameter Estimation and Control Design for Mechanical Systems with Uncertain Parameters*. PhD thesis, Department of Mechanical Engineering, VirginiaTech University, 2010.
- [9] G. Christakos. *Random Field Models in Earth Sciences*. Academic Press, San Diego, CA, 1992.
- [10] J. A. Christen and C. Fox. MCMC using an approximation. *Journal of Computational and Graphical Statistics*, 14(4):795 – 810, 2005.
- [11] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.
- [12] Geir Evensen. The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Systems Magazine*, 29:82–104, 2009.
- [13] D. Gamerman and H. F. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall/CRC, 2006.
- [14] R. Ghanem. Ingredients for a general purpose stochastic finite element implementation. *Computer Methods in Applied Mechanics and Engineering*, 168(1–4):19–34, 1999.
- [15] R. Ghanem. Stochastic finite elements for heterogeneous media with multiple random non-Gaussian properties. *Journal of Engineering Mechanics*, 125:24–40, 1999.
- [16] R. Ghanem and R. Kruger. Numerical solutions of spectral stochastic finite element systems. *Computer Methods in Applied Mechanics and Engineering*, 129(3):289–303, 1996.
- [17] R. Ghanem and P. D. Spanos. *Stochastic finite elements—A spectral approach*. Springer-Verlag, New York, 1991.
- [18] M. Goldstein and D. Wooff. *Bayes linear statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2007.

- [19] S. B. Hazra, H. Class, R. Helmig, and V. Schulz. Forward and inverse problems in modeling of multiphase flow and transport through porous media. *Computational Geosciences*, 8(1):21–47, 2004.
- [20] T. Hida, H.-H. Kuo, J. Potthoff, and L. Streit. *White Noise Analysis-An infinite dimensional calculus*. Kluwer, Dordrecht, 1993.
- [21] D. Higdon, H. Lee, and C. Holloman. Markov chain Monte Carlo-based approaches for inference in computationally intensive inverse problems. *Bayesian Statistics*, (7):181–197, 2003.
- [22] H. Holden, B. Øksendal, J. Ubøe, and T.-S. Zhang. *Stochastic Partial Differential Equations*. Birkhäuser Verlag, Basel, 1996.
- [23] S. Janson. *Gaussian Hilbert spaces*. Cambridge University Press, 1997.
- [24] M. Jardak, C.-H. Su, and G. E. Karniadakis. Spectral polynomial chaos solutions of the stochastic advection equation. In *Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala)*, volume 17, pages 319–338, 2002.
- [25] B. N. Khoromskij and A. Litvinenko. Data sparse computation of the Karhunen-Loève expansion. *Numerical Analysis and Applied Mathematics: International Conference on Numerical Analysis and Applied Mathematics, AIP Conf. Proc.*, 1048(1):311–314, 2008.
- [26] B. N. Khoromskij, A. Litvinenko, and H. G. Matthies. Application of hierarchical matrices for computing Karhunen-Loève expansion. *Journal Computing*, 84(1-2):49–67, 2009.
- [27] T. G. Kolda and B. W. Bader. Matlab tensor classes for fast algorithm prototyping. Technical report, ACM Trans. Math. Software, 2004.
- [28] P. Krée and Ch. Soize. *Mathematics of random phenomena*. D. Reidel Publishing Co., Dordrecht, 1986.
- [29] A. Kučerová and H. G. Matthies. Uncertainty updating in the description of heterogeneous materials. *Technische Mechanik*, 30((1-3)):211–226, 2010.
- [30] O. P. Le Maître, H. N. Najm, R. G. Ghanem, and O. M. Knio. Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *Journal of Computational Physics*, 197(2):502–531, 2004.

- [31] Jia Li and D. Xiu. A generalized polynomial chaos based ensemble Kalman filter with high accuracy. *Journal of Computational Physics*, 228(15):5454–5469, August 2009.
- [32] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley and Sons, Inc., New York, 1969.
- [33] X. Ma and N. Zabaras. An efficient Bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method. *Inverse Problems*, 25(3):035013, 2009.
- [34] P. Malliavin. *Stochastic Analysis*. Springer Verlag, Berlin, 1997.
- [35] Y. M. Marzouk, H. N. Najm, and L. A. Rahn. Stochastic spectral methods for efficient Bayesian solution of inverse problems. *Journal of Computational Physics*, 224(2):560–586, June 2007.
- [36] Y. M. Marzouk and D. Xiu. A stochastic collocation approach to Bayesian inference in inverse problems. *Communications in Computational Physics*, 6(4):826–847, 2009.
- [37] H. G. Matthies. Computational aspects of probability in non-linear mechanics. In A. Ibrahimbegović and B. Brank, editors, *Engineering Structures under Extreme Conditions. Multi-physics and multi-scale computer models in non-linear analysis and optimal design of engineering structures under extreme conditions*, volume 194 of *NATO Science Series III: Computer and System Sciences*. IOS Press, Amsterdam, 2005.
- [38] H. G. Matthies. Stochastic finite elements: Computational approaches to stochastic partial differential equations. *Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM)*, 88(11):849–873, 2008.
- [39] H. G. Matthies, Christoph E. Brenner, Christoph G. Bucher, and Carlos Guedes Soares. Uncertainties in probabilistic numerical analysis of structures and solids—stochastic finite elements. *Structural Safety*, 19(3):283–336, 1997.
- [40] H. G. Matthies and Ch. Bucher. Finite elements for stochastic media problems. *Computer Methods in Applied Mechanics and Engineering*, 168(1–4):3–17, 1999.
- [41] H. G. Matthies and A. Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 194(12-16):1295–1331, 2005.

- [42] H. G. Matthies and E. Zander. Solving stochastic systems with low-rank tensor compression. *Linear Algebra and Application*, submitted, 2010.
- [43] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 26(3):735–747, 2005.
- [44] O. Pajonk, B. V. Rosić, A. Litvinenko, and H. G. Matthies. A deterministic filter for non-Gaussian Bayesian estimation. *Physica D: Nonlinear Phenomena*, 2011.
- [45] B.L. Pence, H.K. Fathy, and J.L. Stein. A maximum likelihood approach to recursive polynomial chaos parameter estimation. In *Proceedings of American Control Conference (ACC)*, pages 2144–2151, 2010.
- [46] L. J. Roman and M. Sarkis. Stochastic Galerkin method for elliptic SPDEs: a white noise approach. *Discrete and Continuous Dynamical Systems. Series B. A Journal Bridging Mathematics and Sciences*, 6(4):941–955, 2006.
- [47] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*. Wellesley-Cambridge Press, Wellesley, MA, 1988.
- [48] A. M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.
- [49] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, 2005.
- [50] A. Tarantola. Popper, Bayes and the inverse problem. *Nature Physics*, 2(8):492–494, August 2006.
- [51] J. Wang and N. Zabaras. Using Bayesian statistics in the estimation of heat source in radiation. *International Journal of Heat and Mass Transfer*, 48(1):15–29, January 2005.
- [52] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60:897–936, 1938.
- [53] D. Xiu and G. E. Karniadakis. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Computer Methods in Applied Mechanics and Engineering*, 191:4927–4948, 2002.
- [54] X. F. Xu. A multiscale stochastic finite element method on elliptic problems involving uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 196(25-28):2723–2736, 2007.

- [55] E. Zander. Stochastic Galerkin library. <http://github.com/ezander/sglib>, 2008.
- [56] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method*. Butterworth-Heinemann, Oxford, 5th ed., 2000.

Technische Universität Braunschweig
Informatik-Berichte ab Nr. 2008-03

2008-03	R. van Glabbeek, U. Goltz, J.-W. Schicke	Symmetric and Asymmetric Asynchronous Interaction
2008-04	R. van Glabbeek, U. Goltz, J.-W. Schicke	On Synchronous and Asynchronous Interaction in Distributed Systems
2008-05	M. V. Cengarle, H. Grönniger B. Rumpe	System Model Semantics of Class Diagrams
2008-06	M. Broy, M. V. Cengarle, H. Grönniger B. Rumpe	Modular Description of a Comprehensive Semantics Model for the UML (Version 2.0)
2008-07	C. Basarke, C. Berger, K. Berger, K. Cornelsen, M. Doering J. Effertz, T. Form, T. Gülke, F. Graefe, P. Hecker, K. Homeier F. Klose, C. Lipski, M. Magnor, J. Morgenroth, T. Nothdurft, S. Ohl, F. Rauskolb, B. Rumpe, W. Schumacher, J. Wille, L. Wolf	2007 DARPA Urban Challenge Team CarOLO - Technical Paper
2008-08	B. Rosic	A Review of the Computational Stochastic Elastoplasticity
2008-09	B. N. Khoromskij, A. Litvinenko, H. G. Matthies	Application of Hierarchical Matrices for Computing the Karhunen-Loeve Expansion
2008-10	M. V. Cengarle, H. Grönniger B. Rumpe	System Model Semantics of Statecharts
2009-01	H. Giese, M. Huhn, U. Nickel, B. Schätz (Herausgeber)	Tagungsband des Dagstuhl-Workshops MBEES: Modellbasierte Entwicklung eingebetteter Systeme V
2009-02	D. Jürgens	Survey on Software Engineering for Scientific Applications: Reuseable Software, Grid Computing and Application
2009-03	O. Pajonk	Overview of System Identification with Focus on Inverse Modeling
2009-04	B. Sun, M. Lochau, P. Huhn, U. Goltz	Parameter Optimization of an Engine Control Unit using Genetic Algorithms
2009-05	A. Rausch, U. Goltz, G. Engels, M. Goedicke, R. Reussner	LaZuSo 2009: 1. Workshop f—r langlebige und zukunfts—fähige Softwaresysteme 2009
2009-06	T. Müller, M. Lochau, S. Detering, F. Saust, H. Garbers, L. Martin, T. Form, U. Goltz	Umsetzung eines modellbasierten durchgängigen Entwicklungsprozesses für AUTOSAR-Systeme mit integrierter Qualitätssicherung
2009-07	M. Huhn, C. Knieke	Semantic Foundation and Validation of Live Activity Diagrams
2010-01	A. Litvinenko and H. G. Matthies	Sparse data formats and efficient numerical methods for uncertainties quantification in numerical aerodynamics
2010-02	D. Grunwald, M. Lochau, E. Börger, U. Goltz	An Abstract State Machine Model for the Generic Java Type System
2010-03	M. Krosche, R. Niekamp	Low-Rank Approximation in Spectral Stochastic Finite Element Method with Solution Space Adaption
2011-01	L. Martin, M. Schatalov, C. Knieke	Entwicklung und Erweiterung einer Werkzeugkette im Kontext von IT-kosystemen
2011-02	B. V. Rosić, A. Litvinenko, O. Pajonk, H. G. Matthies	Direct Bayesian update of polynomial chaos representations